BLOCKING AND TREATMENT EFFECTS AND ORDER OF CONSIDERATION IN STATISTICAL ANALYSES

by

Walter T. Federer

BU-1291-M* May 1995

ABSTRACT

A discussion of blocking or design variables, treatment variables, and model considerations is presented with the idea that confusion arising among these items will be clarified for experimenters and statistical consultants. Spatial trends or gradients and covariates are considered to be blocking factors to be controlled by proper design and / or statistical analysis. An ordering of effects to be considered in a statistical analysis is presented. Basically, design variables and then model considerations must be accounted for prior to obtaining results for the various types of treatment effects.

* In the Technical report Series of the Biometrics Unit, Cornell University, Ithaca, New York 14853

BLOCKING AND TREATMENT EFFECTS AND ORDER OF CONSIDERATION IN STATISTICAL ANALYSES

by
Walter T. Federer

1. INTRODUCTION

In non-orthogonal situations, confusion may arise over the various types of experiment design effects (blocking) and various types of treatment design effects as they relate to their order of consideration in a statistical analysis and to the description of the response model equation used for the statistical analysis as well as the interpretation of the results. Treatment effects have sometimes been treated as blocking effects and *vice versa* by individuals analyzing results from an experiment. In the following, we list types of blocking (non-treatment) and treatment effects and the order(s) that they should be considered in a statistical analysis. Statistical procedures are accompanied by a set of assumptions which need to be fulfilled. Many procedures are based on the assumption of no interaction between the treatment effects and the blocking effects. Interactions with some of the blocking variables can, in some instances, be eliminated by transforming the responses to another scale of measurement and then back transforming the results to the original scale for interpretation. Standardizing responses by dividing by their standard errors may be useful in some cases, especially for variance heterogeneity and unequal numbers of observations. To make clear the distinction between design (blocking) effects and treatment effects, it is necessary to study the types and nature of variation in the experimental material *prior* to applying a treatment to an experimental unit (eu, the unit of experimental material to which one treatment is applied). Then it is necessary to determine how the particular treatments selected (the treatment design) are affected by the various sources of variation present in the experiment. By completely modeling the variation *before and after* applying the treatments to the eus, it is possible to distinguish between blocking effects and treatment effects in the experiment. A modeling of experimental variation necessarily includes a complete description of the population structures for both the population of which the experiment is to be a representative (random) sample for inference purposes and of the population structure used in the experiment if different from the former. Statistical literature appears to be relatively devoid of discussion on this topic except in a few isolated instances (See, e. g., Fisher 1950, pages 2-3, and 1935, chapter II; Federer, 1991, Chapter 7, and 1993, Chapter 10, and references therein.). Except for the preceding, no other references describing population structures for, say, a randomized block designed experiment were found. Present discussions of spatial analyses do little to clarify and many things to muddle the above points, e. g., randomization restrictions due to blocking are ignored for many of the proposed procedures.

In the following, we shall discuss the various types of blocking effects and how they relate to gradients or trends in the experimental material. This is followed by a discussion of the various types of treatment effects. Appropriate design and/or appropriate scales of measurement may be helpful in eliminating certain treatment effects such as environmental interactions and inter-experimental unit competition.

In an experiment, statistical procedures, whether design or analysis, are selected

(i) to obtain a valid estimate of an error variance,

(ii) to reduce error variation by removing extraneous variability from the responses,

and / or

(iii) to obtain a more accurate estimate of the treatment effects.

Methods for doing this have been described in several places in the literature (E. g., see references in Federer, 1994a.). Randomization is used to obtain (i). Blocking, trend analysis, and covariance are useful in accomplishing (ii) and experiment design, trend analysis, covariance, an appropriate response model, and statistical analyses are used to accomplish (iii). In many experimental situations, (i), (ii), and (iii) are the desired goals.

2. POPULATION AND EXPERIMENT STRUCTURES

The goal of an experimenter is to be able to make inferences to some target population using the results obtained from an experiment. If the experiment is a representative or random sample from the target population, then the inferences from the experimental responses to the target population are valid. If the population of which the experiment is a representative sample is not the desired population, inferences from the experimental results to the target population

may be invalid.
Example 1. To illustrate the above, suppose that we are interested in relative sales of v brands of a product, say orange juice, and we conduct an experiment in convenience stores with the v brands simultaneously displayed in an appropriate experiment design. In this experiment, the customer has a choice of v brands from which to select one or more of the brands whereas in practice there may be only one brand available in a store. A customer's buying habits and relative sales may and probably will be completely different when there is choice and when there is no choice of brands. All brands may have the same volume of sales when offered alone but have vastly different volumes of sales when there is a choice of the v brands. In this case, the population structure for the experiment is completely different from the target population where inferences are desired. A useful axiom to follow for experimentation is:

> Axiom 1: The conditions for the experiment must be representative of those in the population for which inferences are to be made. Or, stated another way, the conditions of the experiment must be the same as those used in practice in order to make valid inferences.

Sometimes the population of interest does not include the blocking variables in an experiment. For example, the way plants are handled in a greenhouse prior to transplanting in the field may need to be considered in an experiment but may be irrelevant in practice. Also, it may not be possible to plant an entire experiment in a single day in which case the planting will end at the end of a complete block; this allows the day of planting and complete block to be completely confounded. Several other conditions may necessitate blocking in an experiment. The experimenter can, and often does, include blocking factors which may not be present in the target population. If there are no treatment by blocking factor interactions (a treatment effect), then the inferences from the experiment to the target population are valid, even though the experiment population and the target population differ. It is useful to distinguish among the sampling unit (su, the units making up a population), the observational unit (ou, the unit of observation), and the eu.
Example 2. Given that a teaching method is applied to a class of 30 people, the class is the eu, the individual student is the su, and one of several tests given to a student is the ou. Given an animal feeding experiment, the pen of n animals is the eu, the individual animal is the su, and the weight of the animal at time t is the ou.

3. VARIATION IN THE EXPERIMENT AND BLOCKING

In planning an experiment, a set of experimental material is available to an experimenter. This set is subdivided into N, say, experimental units (eus). There will be variation among the N eus, and the experimenter blocks the eus into blocks in such a manner as to maximize the variation among blocks and to minimize the variation within blocks. Once the blocking is completed, there should be no further way to reduce the variation within blocks that cannot be taken care of through statistical analysis, i.e., there is relative homogeneity among the eus within blocks. If possible, all trends or gradients among the eus within blocks should be absent. If there is a gradient in one direction, this can be eliminated by laying out the eus perpendicular to the gradient (See Federer, 1955, Chapter III.). However, despite an experimenter's best effort to block for sources of variation owing to lack of knowledge or owing to unforeseen events (e.g., bird damage, floods, fires, etc.) that occur during the conduct of the experiment, trends among the eus within a block may occur. The ensuing statistical analysis will need to account for this. Certain experiment designs and statistical analyses proposed by Cox (1958) and Federer (1994b) can be effective in controlling for the effects of differential gradients in experiments.

When an event is unforeseen or happens during the course of the experiment and is not a treatment effect such as water covering a part of the experiment for a period, insect damage to a portion of the experiment, a failure to control weeds in a portion of the experiment, fire damage in a store, etc., covariance for amount of damage or using another block to designate the damaged portion should be used. Of course this would change an orthogonal experiment design into a non-orthogonal one, but with the computing power available today this presents little difficulty for the statistical analyst.

The following axiom may be used to distinguish between blocking variables and treatment variables:

> Axiom 2. All factors affecting variation in experimental material *prior* to the application of treatments are candidates for blocking. Any non-treatment variable that occurs during the course of the experiment is also a candidate for blocking.

This means that factors affecting *all* treatments in the same manner, e.g., an additive effect, may be considered for blocking.

Covariance analysis may be used to eliminate certain types of non-treatment effects affecting responses and here is considered a blocking procedure. Covariance analyses have as their goal the reduction of the error mean square and increased accuracy of the estimated treatment effects. The regression coefficient must then be computed from the error line in an analysis of variance. In more complex experiments there may be several error lines and consequently several regression coefficients (See e.g., Federer and Meredith, 1992). If there are treatment effects for the covariates, then some form of multivariate analysis is required. The statistical analyst must carefully consider co-variables to determine whether they are treatment variables or blocking variables.

Example 2. Snedecor (1946), Table 12.13, presents a data set from a randomized complete block designed experiment for yield of sugar beets with seven fertilizer treatments and with number of beets per eu as a covariate. Since the treatments affected number of beets per eu and not weight per beet (See Steel and Federer, 1955), number of beets is not a suitable covariate as it is a treatment variable.

The various types of blocking variables are:

(i) blocks

complete
incomplete
rows or orders
columns
gradients within blocks
covariates
post-blocking during course of experimentation

The order of consideration in an analysis of the data is complete blocks first, then incomplete blocks and/ or rows and columns, then post-blocking, and then gradients within blocks.

Example 3. Suppose that an experiment was designed as a lattice square and that during the course of the experiment it was noted that there were differential gradients appearing in each of the rows. Cox (1958) has shown how to handle the analysis for a latin square design and Federer (1994b) has given the analysis for lattice rectangle or incomplete block designed experiments and recovering inter-gradient information. In addition, one could fit a second degree polynomial to the rows and to the columns of the lattice rectangle and add interaction terms as Cox and Meeker (1992) did for the latin square and for the lattice rectangle as described by Federer (1994b).

With respect to covariates, these are fitted last since the regression is on the error line(s) of an analysis of variance.

4. EXPERIMENTAL RESPONSES AND TREATMENT EFFECTS

The application of a treatment to an eu may produce a variety of effects. There will be the **direct effect** of a treatment as measured from the responses obtained from the several eus to which the treatment is applied. There may be an **interaction effect** of the treatment with any of a number of environmental factors present in the experiment and/or with other factors in the treatment design. For eus used over several periods, there may be a **carry-over or residual effect** of the treatment used in the preceding periods. Also, if appropriate eus are not used (See Kempton, 1982, and Federer and Basford, 1991.), a treatment may affect all surrounding eus thus exhibiting a **competition effect**. (Note that inter-eu rather than intra-eu competition is the kind of competition referred to here.) To check for competition, one may use the Kempton (1982) single-degree-of-freedom for competition (See Federer, 1994a). Since many experimenters conducting field experiments have used small eus, it is suspected that many of these experiments contain competition effects which have been ignored and it is wondered how many of the results from previous experiments have been vitiated.

In making comparisons among a set of treatments, the experimenter may wish to study effects free of competition, free of residual effects of the previous treatment, and/or on an interaction-free basis. For the last item, a transformation, say a logarithmic transformation, may free the treatment effects from interaction with environmental effects or even other treatment factors. In a repeated measures experiment, direct, residual, or cumulative effects free of the other effects may be the item to be considered. In certain cases, the cumulative effect is the one desired. If so, the experimenter may wish to take repeated measures on the same treatment over a longer time period rather than use shorter periods and changing treatments. The goals of an experimenter will determine the types and order of treatment effects to be used in the analysis. When inter-eu competition effects are present, these effects will need to be removed prior to making comparisons among the treatments. Inter-eu competition arises not from the treatments themselves but from the method used to lay out the experiment and hence is a removable treatment effect and, in many instances, not of interest.

In certain situations, a treatment by environmental interaction may be present and the experimenter wishes to use a parsimonious model with as few parameters as possible. One particular analysis for doing this is the so-called AMMI (additive main effects and multiplicative interaction) model. This analysis can increase the accuracy of the cell means over the full interaction model. As stated above, a transformation of the responses may remove certain types of interaction and make an AMMI analysis unnecessary.

5. MODEL AND SCALE OF MEASUREMENT

Considerable thought should be given to the exact nature of the model equation for responses from an experiment *prior* to adding treatments and *after* adding treatments. *Assuming* that a model given in a statistical textbook is appropriate can be quite inappropriate for the data at hand. The statement that "*the* linear model is" is incorrect at best as all that can be said for most situations is that this is "*a* linear model". In fact, the response model is often non-linear. As William Lawton (personal communication) once remarked, "I never met a linear model in all my work at Eastman Kodak". As a first approximation, the experimenter often uses a linear model to summarize the results from an experiment. A transformation of data to, say logarithms, reciprocals, or square roots, often tends to make a linear model more and the assumptions underlying the statistical analysis more appropriate. Certain types of treatment by environment interactions can sometimes be removed by using data on a transformed scale of measurement, e. g., logarithms. There are situations where an analysis on the obtained responses are needed even though the assumptions for standard statistical analyses are violated. This means that different types of analyses will need to be used when analyses on transformed data do not meet the goal of the experimenter.
Example 4. Given two sets of observations 89.14, 100.00, 112.21 and 10, 100, 1000, the arithmetic means are 100.45 and 370.00, respectively. The arithmetic means of the logarithms of the two sets are both 2 and the back-transformed means are 100. This is not a procedure an experimenter would use for the two data sets.

6. ORDER FOR REMOVING EFFECTS IN THE STATISTICAL ANALYSIS

In performing a statistical analysis on the data from an experiment, the various effects should be taken account of in the order described below. The first item that must taken into account is

scale of measurement and the response model equation
both before and after treatments are applied.

This considers the adequacy of the model and the *additivity* of effects Once this is decided upon, the form, but not the order of effects, of the statistical analysis is determined and set. The first item to list in an analysis such as an analysis of variance (ANOVA) is the block effects in the order listed in Section 3. Then, any trends or gradients within the blocks are taken into account. After this, the treatment effects are considered. If interest centers on the direct effect of treatments, any competition effects or residual effects from previous treatments must be removed to obtain such things as a sum of squares due to direct effects eliminating all other blocking and treatment effects. Since the effect of related variates (covariates) may need to be taken into account, an analysis of covariance (ANCOVA) may need to be performed. Note that the error regression is used to adjust treatment effects for covariate effects. Hence the covariates effects are the last effects to be included in an analysis. This means that the regression due to the covariates *eliminating all other effects,* i. e. the error regression, is the one used. Then, treatment effects are adjusted using the error regression(s) of the covariate(s).

When the blocking effects are random variables and there is partial confounding among the effects, it is possible to recover this information and increase the precision of the estimated treatment effects. Cox (1958) considered the differential gradients in a latin square as fixed effects whereas Federer (1994b) considered the differential gradients in the rows of a lattice square as random effects and recovered inter-gradient information.

In certain situations, the blocks in a block design interact with the treatments and this interaction component of variance is necessarily part of the appropriate error mean square for differences of treatment effects. The commonly assumed response model for a block design is

$$Y_{ij} = \mu + \rho_i + \tau_j + \varepsilon_{ij}, \qquad (1)$$

(where the usual definition of effects is used) whereas it could be

$$Y_{ij} = \mu + \rho_i + \tau_j + \rho\tau_{ij} + \varepsilon_{ij} \qquad (2)$$

or some other more complicated response model. Instead of the error variance being $\sigma_\varepsilon{}^2$ as for equation (1), it is $\sigma_\varepsilon{}^2 + \sigma_{\rho\tau}{}^2$. In order to obtain an unbiased estimate of the error variance, it is necessary to have a random sample of blocks from the population of blocks whereas for the former model and orthogonal experiment designs any sample of blocks would suffice. This fact receives little or no attention in statistical texts, but is in agreement with Fisher's

(1935) definition of a valid error variance for differences of treatment effects in that an appropriate error variance must contain all sources of variation in the difference except that due to the treatments themselves.

When random effects among blocking variables and a partial confounding of treatment and block effects is present in an experiment, various sums of squares eliminating all other effects will need to to computed. This means that several orderings of the random and treatment effects will be used to obtain the desired analysis. The same will be true in testing for significance of effects.

7. LITERATURE CITED

Cox, C. P. (1958). The analysis of Latin square designs with individual curvatures in one direction. Journal of the Royal Statistical Society, Series B, 20:193-204.

Cox, C. P. and J. B. Meeker (1992). A simple alternative to the standard statistical model for the analysis of field experiments with latin square designs. Pre-print Number 92-31, Statistical Laboratory, Iowa State University, Ames, Iowa.

Federer, W. T. (1955). Experimental Design--Theory and Application. Macmillan, New York, pp. xix + 544 + 47. (Republished as the Indian Edition by Oxford & IBH Publishing Company, New Delhi, Calcutta, in 1967).

Federer, W. T. (1991). Statistics and Society--Data Collection and Interpretation, second edition. Marcel Dekker, Inc., New York,, Basel, Hong Kong, pp. xvi + 578.

Federer, W. T. (1993). Statistical Design and Analysis for Intercropping Experiments. Volume I: Two Crops. Springer-Verlag, New York, Heidelberg, London, Paris, Tokyo, Hong Kong, Barcelona, Budapest, pp.xx + 298.

Federer, W. T. (1994a). Diagnostic procedures for analysis of variance. In *New Progress in Probability and Statistics, Proceedings, Conference "Nicolas de Arriquibar", 4th International Meeting of Statistics in the Basque Country, IMSIBAC-4*, VSP-International Science Publishers, Zeist, The Netherlands, "Nicolas de Arriquibar Lecture, pp. 1-23.

Federer, W. T. (1994b). Gradients in incomplete block and resolvable row-column experiment designs. BU-1259-M in the Technical Report Series of the Biometrics Unit, Cornell University, August.

Federer, W. T. and K. E. Basford (1991). Competing effects designs and models for two-dimensional field arrangements. Biometrics 47:1461-1472.

Federer, W. T. and M. P. Meredith (1992). Covariance analysis for split-plot and split-block designs. The American Statistician 46(2):155-162.

Fisher, R. A. (1935). The Design of Experiments. Oliver & Boyd Ltd., Edinburgh (Note: there are several later editions of this text.)

Fisher, R. A. (1950). Statistical Methods for Research Workers. Hafner Publishing Co., New York.

Kempton, R. (1982). Adjustment for competition between varieties in plant breeding trials. Journal of Agricultural Science 98:599-611.

Snedecor, G. W. (1946). Statistical Methods, 4th edition. The Iowa State College Press, Ames, Iowa.

Steel, R. G. D. and W. T. Federer (1955). Yield-stand analyses. Indian Journal of Agricultural Statistics VIII (1-2):27-45.